Research shapes policy: but the dynamics are subtle

Robbie Foy,[1] Louise Locock, [2] Sarah Purdy,[3] Catherine O'Donnell,[4] Nicola Gray,[5] Tim Doran,[6] Huw Davies[7]

[1] Professor Robbie Foy, Leeds Institute of Health Sciences, Charles Thackrah Building, University of Leeds, 101 Clarendon Road, Leeds, LS2 9LJ
r.foy@leeds.ac.uk

Robbie Foy is Professor of Primary Care at the Leeds Institute of Health Sciences, University of Leeds and works one day a week as a general practitioner.

[2] Dr Louise Locock, Department of Primary Health Care Sciences, University of Oxford, 2nd floor, 23-28 Hythe Bridge Street, Oxford OX1 2ET
louise.locock@phc.ox.ac.uk

Louise Locock is Deputy Research Director in the DIPEx Health Experiences Research Group and a University Research Lecturer in the Department of Primary Health Care Sciences, University of Oxford.

[3] Dr Sarah Purdy, Academic Unit of Primary Health Care, University of Bristol, 39 Whatley Road, Clifton, Bristol BS8 2PS
sarah.purdy@bristol.ac.uk

Sarah Purdy is a Consultant Senior Lecturer/MRC Clinician Scientist in the Academic Unit of Primary Health Care, University of Bristol and also works as a general practitioner.

[4] Professor Catherine O'Donnell, Institute of Health & Wellbeing, University of Glasgow, 1 Horselethill Road, Glasgow G12 9LX
Kate.O'Donnell@glasgow.ac.uk

Catherine O'Donnell is Professor of Primary Care Research and Development at the Institute of Health & Wellbeing, University of Glasgow.

[5] Dr Nicola Gray, Independent Researcher, Manchester, Lancashire
nicola@webstar.co.uk

Nicola Gray is an Independent Pharmacist Researcher based in Manchester, Lancashire

[6] Dr Tim Doran, Clinical Research Fellow in Public Health, National Primary Care Research & Development Centre, University of Manchester, Williamson Building, Oxford Road, Manchester M13 9PL
Tim.Doran@manchester.ac.uk

Tim Doran is Clinical Research Fellow in Public Health at the National Primary Care Research and Development Centre, University of Manchester.

[7] Professor Huw T. O. Davies, Professor of Health Care Policy and Management, School of Management, University of St Andrews, The Gateway, North Haugh, St Andrews, Fife KY16 9RJ
hd@st-andrews.ac.uk

Huw Davies is Professor of Health Care Policy and Management and Co-head of School at the School of Management, University of St Andrews, and Co-Director of the Social Dimensions of Health Institute, Universities of Dundee and St Andrews.

Corresponding author:

Professor Huw Davies
School of Management
University of St Andrews
The Gateway
North Haugh
St Andrews
Fife
KY16 9RJ

Email: hd@st-and.ac.uk
Tel: +44 (0) 1334 462870

Main text: 2624 words

Summary

Major policy initiatives such as the Quality and Outcomes Framework (QOF) in the national contract for general practitioners might variably be informed by evidence at their inception, implementation and subsequent evolution. But what evidence gets admitted into these policy debates – and what is left out? Using QOF as an example, this article demonstrates what an analysis of the relationship between policy and the associated research can tell us about the underlying policy assumptions and about the role of evidence in policy debates.

Introducing the Quality and Outcomes Framework (QOF)

The formation and implementation of policy are not especially linear, wholly rational, or always fully evidence-based (Weiss, 1977; Weiss, 1982). Understanding these dynamics, and teasing out the role of research evidence, has been a major preoccupation of researchers and policy analysts in recent years – with the hope that a better understanding of these processes can lead to better policy making and improved services. While policy typically changes only incrementally, sometimes circumstances conspire to prompt more radical shifts (Baumgartner and Jones, 1993). One such radical shift occurred in the United Kingdom in 2004 with the incorporation of the Quality and Outcomes Framework (QOF) into a new national contract for general practitioners (GPs), negotiated between the Department of Health (DH) and the British Medical Association (BMA) (Roland, 2004). The new contract aimed to increase GP rewards and improve the quality of patient care by linking substantial payments to practice performance on 146 publicly-reported quality indicators.

Several factors aligned to create the necessary conditions for the introduction of quality-related financial incentives. First, pay for GPs had fallen behind that of specialists working in hospitals, and general practice was suffering from low morale and a recruitment and retention crisis (British Medical Association, 2001). Second, evidence had emerged that quality of care in many settings was highly variable (Institute of Medicine, 2001; Leatherman and Sutherland, 2003). Third, technical improvements now made it possible to measure quality of care more accurately, and several countries had begun experimenting with pay-for-performance systems derived from industry (Mannion

and Davies, 2008). Finally, the UK government had committed to improving quality and equity of care, supported by additional investment in the NHS. From this convergence of factors, a consensus emerged between the profession and the DH that a new GP contract was required that would incentivise doctors to work in general practice and improve access to high quality care (National Audit Office, 2008). The QOF was the centrepiece of these new arrangements (Shekelle 2003; Roland, 2004).

Since its inception there has been debate about how the QOF should be maintained and developed, and it is subject to annual review. Potential revisions are supported by a review of relevant evidence (more recently, by The National Institute for Health and Clinical Excellence, NICE 2011), but the final decisions on changes to the framework result from contract negotiations between the DH and the BMA. This article, however, does not focus on the QOF *per se*. Instead, we see what an analysis of the relationship between the QOF and associated research can tell us about underlying policy assumptions and the role of evidence in shaping policy debates.

Research and policy

Various models have been proposed to explain policy developments. Models that emphasise the chaotic nature of policy making, such as the 'garbage can' model (Cohen *et al.*, 1972), regard the attachment of solutions to problems as a largely arbitrary or chance process. More structured views of the policy world see change as being driven by interactions between distinct streams of activity and debate, such as the political and media swirl that highlights particular issues as problems, or the policy and research arenas that identify strategies and tools as potential solutions (Kingdon, 1984). When these streams come together, a major change may be possible – Kingdon identifies such moments as 'policy windows', an opportunity for new ideas to break through and command attention and action. Once such change has happened, a period of incremental modification and consolidation is then likely – as we have seen with the QOF since its implementation. Importantly, however, the particular circumstances that led to the original policy shift will continue to shape the sorts of incremental change possible thereafter (Pierson, 2000).

While by no means the overriding factor, research evidence can play a role at all stages, both shaping and constraining not just the initial policy shift, but also influencing subsequent developments. But why does some evidence come to the fore, whilst other evidence remains unconsidered? And what does this tell us about the interplay between research and policy?

There are generally three types of stances adopted by researchers in influencing public policy: the consensual approach; the contentious approach; and the paradigm-challenging approach (Adapted from Rein, 1976; Weiss, 1995; Nutley *et al.* 2007). The *consensual* approach refers to situations where there is broad agreement among policy makers and researchers about the main issues of concern and the ways in which these should be addressed. Researchers then work within the existing paradigm, aiming to provide policy makers, practitioners and other stakeholders with knowledge about how best to make technical adjustments to improve service delivery and outcomes. In the *contentious* approach, researchers place themselves more on the sidelines of public policy. They may not always contribute to policy development directly, but maintain a more critical stance. In the *paradigm-challenging* approach, researchers might take a stance outside the orthodoxy, working to problematize established ways of thinking.

The research evidence that is most easily considered is that which works with the grain of current policy and preferred solutions, and that essentially offers technical refinements to the existing arrangements – the consensual approach. In that sense, research that offers little fundamental challenge is most easily assimilated into the discussion around policy. More challenging research that calls into question the basic assumptions behind current policy directions is less likely to get a hearing. Taken-for-granted assumptions that underpin established policy directions are more resilient to challenge, often needing new policy leadership before they can be called into question.

This paper then explores the array of research potentially relevant to the QOF. It examines what research evidence has been produced and is having some impact, what evidence has been produced but as yet is failing to influence, and what areas of potential interest and concern are relatively under-explored. We draw upon deliberations from a series of workshops which explored the relationship between evidence and health policy. Workshop participants included primary care academics, with

briefing sessions delivered by policy experts and researchers. After selecting the QOF as a focus to explore the policy/evidence nexus, we drew on the results of structured reviews of the literature on pay-for-performance in healthcare conducted by workshop participants, as well as on other published review sources that focused specifically on QOF. From this account emerges an analysis of *why* certain patterns of research engagement are observed, and this analysis in turn highlights some implications for policy evolution and the potential role of research.

The research discourse on the QOF so far

While it might be tempting to talk of 'bodies of evidence' relating to a particular policy, it makes more sense to talk instead of a research 'discourse' – a way of conversing, and even thinking, about a topic that has clear limits as to what is seen as relevant and appropriate. Figure 1 provides an overview of the research discourse in relation to the QOF to date. It highlights, for example, the focus on quantitative analyses of QOF impact on incentivized care: in general; for specific conditions (particularly coronary heart disease and diabetes); and for underserved populations (mainly socio-economically deprived populations but also minority ethnic groups and the elderly). This analysis has largely been consensus-building in nature, initially reinforcing the assumption that the QOF had improved care for incentivized conditions (e.g. Campbell *et al.*, 2007; Doran *et al.*, 2008). Over time, however, some have questioned the overall value of the QOF, often citing the same literature that had previously been used as evidence of its effectiveness (Gillam, 2010; Siriwardena, 2010). Research that is potentially more contentious, such as that showing that the QOF has had little impact on outcomes, disease management or hospital admissions (Purdy *et al*., 2011; Serumaga *et al.*, 2011; Smith, 2011) has, until recently, generated little discussion or comment.

Other areas of research have been conspicuous by their absence from discussions about the value of the QOF (Figure 1). For example, several predominantly qualitative studies have explored the impact of the QOF on practice nurses (Checkland *et al.*, 2007; McDonald *et al*., 2007; McGregor *et al*., 2008*)* – yet there is little policy-related discussion of negative impacts on practice nurse workload or morale. Similarly, the research discourse on the patient experience has been relatively limited until recently (Hankins *et al.*, 2007), with in-depth or longitudinal research only just starting to emerge (Dowrick *et*

al., 2009; Campbell *et al*., 2010; Coventry *et al*., 2011). Also neglected is an examination of whether the QOF has driven changes in the organization of care that especially affect patients with co-morbidities – requiring them to attend multiple appointments for chronic disease management, which runs counter to calls for 'minimally disruptive medicine' (May *et al.*, 2009). Finally, there is little work exploring the types of governance and regulation that primary care organizations may be employing to monitor and influence practice performance through analysis of QOF data (O'Donnell *et al.*, 2011).

We would suggest that the lack of research discourse in some of these areas is due not only to shortages of data and technical difficulties. It is also a function of the inherently contentious nature of such work and its potential to challenge current thinking about the utility of the QOF. Thus while the QOF itself represented a major shift in physician reimbursement (Shekelle, 2003), we question what evidence is now required to reform or promote further radical shifts of the new orthodoxy.

How can we explain the (im)balance of the research discourse?

The patterning of what gets researched, and what does not – and what has influence, and what does not – is not arbitrary. There are deeper drivers, foremost of which is the original framing of the policy 'problem' that the intervention was designed to address.

The main aims of the QOF were to improve remuneration for GPs, thereby promoting recruitment and retention, and to incentivize high quality care (NHS Confederation and British Medical Association, 2003). These aims were to be met by financially rewarding participation in an annual quality improvement cycle and the achievement of centrally-determined quality targets. Subsequent re-interpretations resulted in research that focused mostly on the technical aspects of the QOF as a pay-for-performance scheme and its capacity to improve quality as defined under the framework. Subsequently, much of the discourse has treated GPs as agents whose behaviour needs to be monitored and controlled in order to prevent abuses of trust. Whilst this is a legitimate line to pursue, it tends to sideline other research on quality improvement that focuses on improving professionalism, wider aspects of the organization of care and the patient experience.

This dominant discourse may also fail to capture some of the less frequently voiced rational or benevolent reasons why GPs might game or circumvent the system: for example, GPs serving more socio-economically deprived populations may see exception coding in QOF as a legitimate means to correct what they see as inequitable resource distribution. Thus the original policy framing, and its subsequent co-option to newer policy discourses, has both enabled and constrained the parallel research discourse.

Placing pay-for-performance centre-stage in this way also smuggles into the policy debate a set of assumptions that require rather more critical re-examination (Mannion and Davies, 2008). For example, pay-for-performance approaches assume that people act rationally in seeking to maximise income. Whilst this may be at least partly true, there is a risk that it fails to capture a wider range of determinants of behaviour. Hence, perceived peer pressure, reputational risk, a desire to conform and the need for approval may drive improvements in the delivery of care beyond upper payment thresholds for the QOF. There is also limited understanding of some of the finer-grained but nonetheless critical aspects of pay-for-performance – such as which organizational or team levels the scheme should operate at, and at what levels are incentives most cost-effective in achieving a desired change in clinical performance (Mannion and Davies, 2008).

As so often with new policy, roll out occurred before evaluation. Thus, the context in which the QOF was negotiated precluded any prior testing of effects. Researchers, constrained by this, could subsequently only play "catch-up" by undertaking quasi-experimental evaluations, which are prone to bias, difficult to interpret and hence provide questionable evidence on effectiveness (Campbell *et al*., 2007; Campbell *et al.,* 2009; Serumaga *et al.*, 2011). Research supply-side issues also provide some explanation for the patterning of the research discourse around the QOF. Earlier research on specific quantitative issues around impact was facilitated by core DH funding and drew upon readily available data generated by the payment system. In contrast, for example, qualitative research challenging some of the assumptions about the mechanisms and benefits of the QOF took longer to emerge, partly because of the time needed to acquire the necessary resources to establish studies and to collect and analyze the data.

The dominant research discourse on the QOF was therefore driven by assumptions about policy objectives and mechanisms, as well as by the context of its implementation. As a result the discourse largely reflects the joint agenda set by GPs and government, each committed to a major new initiative. Work reflecting the interests of other, less powerful, groups, such as primary care nurses and patients, has been relatively slower to emerge and has struggled to get onto the policy agenda. Even now, much of the debate about the future direction of QOF concerns which indicators should be retired or introduced (Anekwe, 2011; NICE, 2011), rather than the more contentious issues about its mechanisms, effects and wider consequences.

Conclusions

We have used the example of the QOF in this paper due to its currency and policy relevance. This discussion is about understanding the development of an evidence base and the subsequent discourse around it, and not about judging the quality or strength of that evidence. Our analysis contends that what evidence gets to the policy table reflects the interests and preconceptions of powerful stakeholder groups and the biases introduced by the differential availability of data. In Kingdon's terms of 'problem', 'policy' and 'political' streams (Kingdon 1984), the 'problem' of needing better remuneration for GPs met with a 'policy stream' which favoured quantitatively measurable targets and financial incentives, and a 'political' sense of mistrust of GPs as autonomous self-employed agents who needed to be controlled.

The evidence base is inevitably incomplete and imperfect – it is therefore important to consider where gaps in discourse lie and what these tell us about the prevailing assumptions and beliefs that influence policy evolution. For the QOF, some of these gaps are extensive. In addition, there has only been a limited response to emerging evidence, with most activity focusing on revision or inclusion of new clinical conditions: so selectivity in the attention paid to even available evidence is also clear.

The example of the QOF provides a timely reminder of the risk of an incomplete body of research informing the policy discourse (availability bias), and draws upon the different typographies of research evidence (consensual, contentious or paradigm-challenging) to illustrate how not all

research is seen as equally relevant and/or influential. In this case, monitoring the success of current indicators, or collating evidence for the insertion of new indicators, seems to be the research that is most readily taken into account. In contrast, research that undermines beliefs about the power of financial incentives, or focuses on the dysfunctional consequences of policy implementation, appears less welcome in debates around the QOF.

There are lessons from this example for other policy developments, such as new commissioning arrangements in health care, changes to the benefits system and the prospects for criminal justice system reform. Above all, this analysis cautions against the legitimizing rhetoric of evidence: when policy advocates make claims that policies are informed by evidence we should ask not only which evidence has shaped the policies, but also what of the evidence that has not.

Acknowledgements

We are grateful to Martin Roland for his comments on an earlier version of this paper. The workshops were supported by the National Institute for Health Research School for Primary Care Research. The views represented here are solely those of the authors.

References

Anekwe, L. (2011), QOF depression indicators face axe. *Pulse*, *9 June.*

Baumgartner, F. and Jones, B.D. (1993), *Agendas and instability in American politics* (University of Chicago Press, Chicago).

British Medical Association (2001), *National survey of GP opinion* (BMA, London).

Campbell, S., Reeves, D., Kontopantelis, E., Middleton, E., Sibbald, B., Roland, M. (2007), Quality of primary care in England with the introduction of pay for performance. *New England Journal of Medicine, 357,* pp.181-90.

Campbell, S.M., Reeves, D., Kontopantelis, E., Sibbald, B., Roland M. (2009), Effects of pay for performance on the quality of primary care in England. *New England Journal of Medicine, 361,* pp.368-78.

Campbell, S.M., Kontopantelis, E., Reeves, D., Valderas, J.M., Gaehl, E., Small, N. *et al*. (2010), Changes in patient experiences of primary care during health service reforms in England between 2003 and 2007. *Annals of Family Medicine*, *8*, pp.499-506.

Checkland, K., McDonald, R., Harrison, S. (2007), Ticking boxes and changing the social world: Data collection and the new UK general practice contract. *Social Policy and Administration, 41,* pp.693-710.

Cohen, M.D., March, J.G., Olsen, J.P. (1972), A garbage can model of organizational choice. *Administrative Science Quarterly*, *17,* pp.1-25.

Coventry, P.A., Hays, R., Dickens, C., Bundy, C., Garrett, C., Cherrington, A. *et al*. (2011), Talking about depression: a qualitative study of barriers to managing depression in people with long term conditions in primary care. *BMC Family Practice*, *12,10*.

Doran, T., Fullwood, C., Kontopantelis, E., Reeves, D. (2008), Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework. *The Lancet*, *372*, pp. 728-36.

Dowrick, C., Leydon, G.M., McBride, A., Howe, A., Burgess, H., Clarke, P. *et al*. (2009), Patients' and doctors' views on depression severity questionnaires incentivised in UK quality and outcomes framework: qualitative study. *British Medical Journal, 338*, b663.

Gillam, S. (2010), Should the Quality and Outcomes Framework be abolished? Yes. *British Medical Journal, 340,* c2710.

Hankins, M., Fraser, A., Hodson, A., Hooley, C., Smith, H. (2007), Measuring patient satisfaction for the Quality and Outcomes Framework. *British Journal of General Practice, 57*, pp.737-40.

Institute of Medicine (2001), *Crossing the quality chasm: A new health system for the 21st century* (National Academy Press, Washington D.C).

Kingdon, J.W. (1984), *Agendas, alternatives and public policies* (Little, Brown, Boston).

Leatherman, S., and Sutherland, K. (2003), *The quest for quality in the NHS* (The Nuffield Trust, London).

Mannion, R., and Davies, H.T.O. (2008), Payment for performance in health care. *British Medical Journal, 336,* pp.306-8.

May, C., Montori, V.M., Mair, F.S. (2009), We need minimally disruptive medicine. *British Medical Journal, 339,* pp.485-7.

McDonald, R., Harrison, S., Checkland, K., Campbell, S.M., Roland, M. (2007), Impact of financial incentives on clinical autonomy and internal motivation in primary care: ethnographic study. *British Medical Journal, 334,* 1357.

McGregor, W., Jabareen H, O'Donnell, C.A., Mercer, S.W., Watt, G.C.M. (2008), Impact of the 2004 GMS contract on practice nurses: a qualitative study. *British Journal of General Practice, 58,* pp. 711-9.

National Audit Office (2008), *NHS Pay Modernisation: New contracts for general practice services in England* (The Stationery Office, London).

National Institute for Health and Clinical Excellence (NICE) (2011), *QOF Advisory Committee recommendations for indicator development for potential inclusion on the NICE menu for the 2012/13 QOF* (NICE, London).

NHS Confederation and British Medical Association (2003), *New GMS contract 2003. Investing in general practice* (NHS Confederation and British Medical Association, London).

Nutley, S.M., Walter, I., Davies, H.T.O. (2007), *Using evidence. How research can inform public services* (Policy Press, Bristol).

O'Donnell, C.A., Ring, A., McLean, G., Grant, S., Guthrie, B., Gabbay, M., Mair, F. S., Carlisle, C., Heaney, D., Sutton, M., and Watt, G. C. M. (2011), *The new GMS contract in primary care: The impact of governance and incentives on care* (NIHR Service Delivery & Organisation Programme, London).

Pierson, P. (2000), Path dependence, increasing returns and the study of politics. *American Political Science Review*, *94*, pp.251-67.

Purdy, S., Griffin, T., Salisbury, C., Sharp, D. (2011), Emergency admissions for coronary heart disease: A cross-sectional study of general practice, population and hospital factors in England. *Public Health, 125,* pp.46-54.

Rein, L. (1976), *Social science and public policy* (Penguin Education, Harmondsworth).

Roland, M. (2004), Linking physicians' pay to the quality of care - A major experiment in the United Kingdom. *New England Journal of Medicine, 351,* pp. 1448-54.

Serumaga, B., Ross-Degnan, D., Avery, A.J., Elliott, R.A., Majumdar, S.R., Zhang, F. *et al*. (2011), Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *British Medical Journal, 342.*

Shekelle, P. (2003), New contract for general practitioners. *British Medical Journal, 326,* pp.457-8.

Siriwardena, N. (2010), Should the Quality and Outcomes Framework be abolished? No. *British Medical Journal, 340,* c2794.

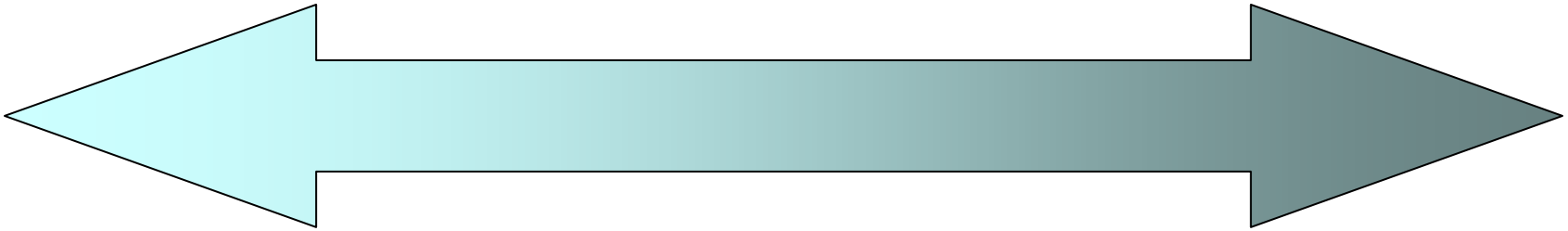
Smith, P. (2011), *Do quality improvements in primary care reduce secondary care costs? Primary research into the impact of the Quality and Outcomes Framework on hospital costs and mortality* (The Health Foundation, London).

Weiss, C.H. (1977), Research for policy's sake: The enlightenment function of social research. *Policy Analysis, 3,* pp. 531-45.

Weiss, C.H. (1982), Policy research in the context of diffuse decision making. *Journal of Higher Education, 53,* pp. 619-39.

Weiss, C.H. (1995), The haphazard connection: Social science and public policy. *International Journal of Educational Research*, *23,* pp. 137-50.

Figure 1. Research discourse in relation to the Quality and Outcomes Framework



Areas with less research attention

Patient perspectives

Impact on patients with multiple morbidities

Models of personal motivation

De-professionalization and values

Impact on practice culture

Decision-making with respect to exception reporting

Impact on non-incentivized conditions

Gaming

Skill mix and impact on professional groups

Practice structure

Exception reporting

Areas with more research attention

Process measures

Indicator mix

Performance versus recording

Impact of QOF on outcomes e.g. admissions

Impact of QOF on inequalities